Nonmonotone VIP : A Bundle Type Approach

Dominique Fortin, Ider Tsevendorj

INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET EN AUTOMATIQUE

Nonmonotone VIP : a bundle type approach

Dominique Fortin, Ider Tsevendorj

N° 4062

Novembre 2000

THÈME 4

Rapport de recherche

ISSN 0249-6399 ISRN INRIA/RR--4062--FR+ENG

INRIA
ROCQUENCOURT

Nonmonotone VIP : a bundle type approach

Dominique Fortin, Ider Tsevendorj

Thème 4 — Simulation et optimisation
de systèmes complexes
Projet Adopt

Rapport de recherche n ° 4062 — Novembre 2000 — 32 pages

Abstract: In this paper we provide an algorithm for solving piecewise convex maximization problems. It occurs as a tool for solving nonmonotone variational inequalities in connection with mathematical programs with equilibrium constraints. We report first computational experiments on small examples and open some issues to improve optimality conditions checking.

Key-words: global optimization, piecewise convex function, nonmonotone variational inequality, equilibrium constraints

(Résumé : tsvp)

Unité de recherche INRIA Rocquencourt
Domaine de Voluceau, Rocquencourt, BP 105, 78153 LE CHESNAY Cedex (France)
Téléphone : 01 39 63 55 11 - International : +33 1 39 63 55 11
Télécopie : (33) 01 39 63 53 30 - International : +33 1 39 63 53 30

Inégalités variationnelles non monotones: une approche par faisceaux

Résumé : Dans ce rapport, nous donnons un algorithme pour résoudre des problèmes de maximisation convexe par morceaux. Il intervient comme outil de base dans la résolution d'inégalités variationnelles non monotones en relation avec des programmes mathématiques soumis à des contraintes d'équilibre. Nous rendons compte des résultats expérimentaux sur de petits exemples et ouvrons quelques enjeux concernant la vérification des conditions d'optimalité.

Mots-clé : optimisation globale, convexe par morceaux, inégalités variationnelles, contraintes d'équilibre

1. INTRODUCTION

A mathematical program with equilibrium constraints is a constrained optimization problem, where some constraints are variational inequalities.

$$\begin{cases} \text{minimize} & \varphi(x,y) \\ \text{subject to} & x \in \Omega, \\ & y \in S(x), \end{cases}$$

where $\varphi : \mathbb{R}^{\ell+m} \to \mathbb{R}$, $\Omega \subset \mathbb{R}^{\ell}$ and $S(x)$ is the solution set of a variational inequality problem:

$$\begin{cases} \text{find} & y \in C(x) \\ \text{such that} & \langle \Phi(x,y), u-y \rangle \geq 0, \quad \text{for all } u \in C(x), \end{cases}$$

for an operator Φ and a set-valued map C. In this paper, we attempt to tackle above nested difficulties from the inner variational constraints viewpoint. We borrow from bundle method, the concept of piecewise approximation of the domain and propose a nonlinear (indeed quadratic) extension; the main contribution is the design and implementation of an algorithm to solve a piecewise quadratic convex maximization problem.

We provide motivations, in section 2, for considering mathematical program with equilibrium constraints as mainly involved with variational inequalities; then, in section 3, we address non monotone variational inequalities and give a first algorithm that converges under a condition to be satisfied; since it is hard to characterize operator fulfilling this condition, we provide in section 4 an alternate piecewise convex maximization approach to solve nonmonotone variational inequalities. First experiments, on small examples in dimension 2 and 3, happen to be effective, however we give some hints how to handle structural properties of intersection of level sets to carry algorithm over higher dimensional space in an efficient way.

2. FROM (MPEC) TO (VIP)

Let an operator $\Phi : \mathbb{R}^{\ell+m} \to \mathbb{R}^m$ and a set-valued map $C : \mathbb{R}^\ell \to \mathbb{R}^m$ be given. Then variational inequality problem is defined as follows

$$\begin{cases} \text{find} & y \in C(x) \\ \text{such that} & \langle \Phi(x,y), u-y \rangle \geq 0, \quad \text{for all } u \in C(x), \end{cases} \tag{1}$$

and its solution set is denoted by $S(x)$.

A mathematical program with equilibrium constraints (MPEC) is a constrained optimization problem, where some constraints are variational inequalities. We consider a (MPEC) problem in the following statement

$$\begin{cases} \text{minimize} & \varphi(x,y) \\ \text{subject to} & x \in \Omega, \\ & y \in S(x), \end{cases} \tag{MPEC}$$

where $\varphi : \mathbb{R}^{\ell+m} \to \mathbb{R}$, $\Omega \subset \mathbb{R}^\ell$ and $S(x)$ is the solution set of the variational inequality problem (1).

From this statement, it is clear that a bilevel optimization problem (BP) is a particular case of (MPEC) since the optimality conditions for lower level problem of (BP) are particular cases of the inclusion type constraint $y \in S(x)$ in $(MPEC)$.

Since the traffic equilibrium problem could be formulated as a variational inequality, many transportation planning and design problems are (MPEC). For a solution of such problems different approaches could be used depending on properties of the set $S(x)$. In fact, properties of $S(x)$ are mostly depending on the operator $\Phi(\cdot,\cdot)$. Therefore we will recall some definitions and proposition from [LR96]. To simplify notation, we fix variable x and give definitions w.r.t. second variable y.

Definition 2.1. An operator $\Phi : \mathbb{R}^m \to \mathbb{R}^m$ is said to be

- monotone on C if for all $u, v \in C$

$$\langle \Phi(u) - \Phi(v), u-v \rangle \geq 0;$$

- strictly monotone on C if for all $u, v \in C$

$$\langle \Phi(u) - \Phi(v), u-v \rangle > 0;$$

- strongly monotone on C if there exists a constant κ, such that for all $u, v \in C$

$$\langle \Phi(u) - \Phi(v), u-v \rangle \geq \kappa \parallel u-v \parallel;$$

• antimonotone on C if $-\Phi(\cdot)$ is monotone; (in other words if for all $u, v \in C$

$$\langle \Phi(u) - \Phi(v), u - v \rangle \leq 0);$$

• nonmonotone on C if it is not monotone on C; (*notice that antimonotone operator is nonmonotone too*).

PROPOSITION 2.1. *[] Let C be a closed convex set in $\mathbb{R}^m$ and $\Phi(\cdot)$ be a continuous mapping. Let S denote the (possibly empty) solution set of the (1).*

•*If $\Phi(\cdot)$ is monotone on C, then S, if nonempty, is a closed convex set.*
•*If $\Phi(\cdot)$ is strictly monotone on C, then S consists of at most one element.*
•*If $\Phi(\cdot)$ is strongly monotone on C, then S consists of exactly one element.*

Due to this proposition, for example, if the operator $\Phi(\cdot)$ is strongly monotone or strictly monotone and $S(x) \neq \emptyset$ (in other words $S(x)$ is singleton) then $(MPEC)$ becomes a standard optimization problem:

$$\begin{array}{rl} \text{minimize} & \varphi(x, S(x)) \\ \text{subject to} & x \in \Omega. \end{array}$$

Also, if $S(x) = \{y \in \mathbb{R}^m \mid \Phi(y) = 0\}$ then $(MPEC)$ becomes an optimization problem with equality type constraints

$$\begin{array}{rl} \text{minimize} & \varphi(x, y) \\ \text{subject to} & x \in \Omega, \\ & \Phi(y) = 0. \end{array}$$

But, unfortunately, such nice cases happen scarcely in practical optimization, in particular in traffic assignment problem operator $\Phi(\cdot)$ is nonmonotone. In the case of nonmonotonocity of $\Phi(\cdot)$ and $S(x) \neq \{y \mid \Phi(y) = 0\}$, it seems to us that the main difficulty to solve $(MPEC)$ highly depends on its (VI) constraint. On the other hand, first order optimality conditions in x for $(MPEC)$ turns out to be a (VI) too, then in an attempt to iteratively solve $(MPEC)$, this paper will be devoted to the question of solving nonmonotone (VIP).

To underline an idea to solve $(MPEC)$ through (VIP) we will assume all necessary good properties on the function $\varphi(\cdot, \cdot)$, sets Ω and $C(x)$ but the monotonicity of the operator. For example, we will assume that function $\varphi(\cdot, \cdot)$ is convex in both variables,operator $\Phi(\cdot)$ does not depend on x etc... .

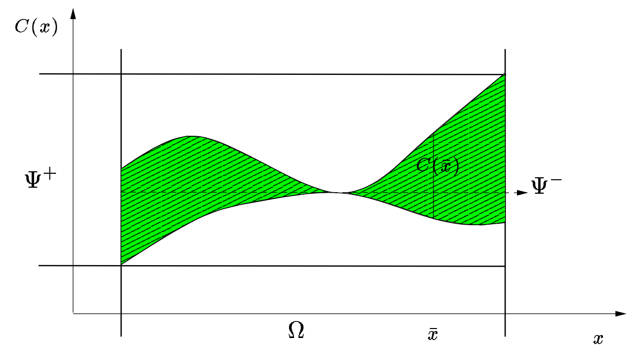

FIG. 1. (MPEC) domain: inner and outer approximation

1. If one could construct an inner approximation Ψ^- such that $\Psi^- \subset C(x)$ for all $x \in \Omega$ and $\Psi^- \neq \emptyset$ then

$$\begin{cases} \langle \nabla_x \varphi(\overline{x}, \overline{y}), x - \overline{x} \rangle \geq 0 \quad \text{for all } x \in \Omega \\ \langle \Phi(\overline{y}), y - \overline{y} \rangle \geq 0, \quad \text{for all } y \in \Psi^- \end{cases} \tag{2}$$

2. If one could construct an outer approximation Ψ^+ such that $C(x) \subset \Psi^+$ for all $x \in \Omega$ then

$$\begin{cases} \langle \nabla_x \varphi(\overline{x}, \overline{y}), x - \overline{x} \rangle \geq 0 \quad \text{for all } x \in \Omega \\ \langle \Phi(\overline{y}), y - \overline{y} \rangle \geq 0, \quad \text{for all } y \in \Psi^+ \end{cases} \tag{3}$$

3. If one could find a subset $P \in \mathbb{R}^m$ such that $S(x) \simeq P$ then denoting

$$\begin{aligned} & z = (x, y) \in \mathbb{R}^{\ell+m}; \\ & \phi(z) = \varphi(x, y); \\ & D = \Omega \times P; \end{aligned}$$

we have the following problem

$$\begin{aligned} \text{minimize} \quad & \phi(z) \\ \text{subject to} \quad & z \in D. \end{aligned}$$

Here, for instance, previous approximations P^+ and P^-, where

$$\begin{aligned} P^+ &= \{y \mid \langle \Phi(y), u - y \rangle \geq 0, \quad \text{for all } u \in \Psi^-\}, \\ P^- &= \{y \mid \langle \Phi(y), u - y \rangle \geq 0, \quad \text{for all } u \in \Psi^+\}. \end{aligned}$$

could be used as candidates for such a P since $S(x) \subset P^+$ and $P^- \subset S(x)$.

for all three different cases, we have to solve:

$$\text{find } \overline{z} \in D: \ \langle T(\overline{z}), z - \overline{z} \rangle \leq 0 \quad \text{for all } z \in D, \tag{4}$$

where $z = (x, y) \in \mathbb{R}^{\ell+m}$,

$$T(z) \ = \ \begin{cases} [-\nabla_x \varphi(x,y), -\Phi(y)]; \text{ in case 1 and 2.} \\ -\nabla \phi(z); \text{ in case 3.} \end{cases}$$

$$D \ = \ \begin{cases} \Omega \times \Psi^-; \quad \text{in case 1.} \\ \Omega \times \Psi^+; \quad \text{in case 2.} \\ \Omega \times P. \quad \text{in case 3.} \end{cases}$$

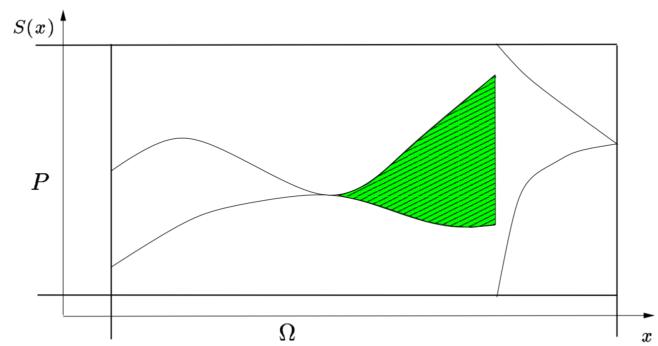

FIG. 2. (MPEC) classification

Notice that (2) (or (3)) implies (4), but the converse is not true in general. So, we will have some lower or upper bounds for $(MPEC)$ by solving (4) which is nonmonotone in general.

In an attempt to explain the reverted sign in (4) compared with (1), we would say that (1) comes from convex minimization while we address in (4) a problem more involved with convex maximization. Notice that (4) is a nonmonotone variational inequality problem even if operator $T(\cdot)$ is a monotone operator.

Since, in the remainder of this paper, we only consider nonmonotone (VIP), we retain the unusual sign $'' \leq''$; however, we follow standard notation to some extent by setting $n = \ell + m$ and $T : \mathbb{R}^n \rightarrow \mathbb{R}^n, \ \ D \subset \mathbb{R}^n$:

$$\text{find } z \in D: \ \langle T(z), x - z \rangle \leq 0 \quad \text{for all } x \in D, \tag{VIP(T,D)}$$

Lemma 2.1. *Let $T(\cdot)$ be any operator, $T(x) \neq 0$ for all $x \in D$; D be a convex and full dimensional subset in $\mathbb{R}^n$. Then if z solves (VIP(T,D)) then $z \in bd(D)$ (boundary point).*

Proof. By contradiction let $z \in D$ solve (VIP(T,D)), $T(z) \neq 0$ and $z \notin bd(D)$. Then, there exists $\varepsilon > 0$ s.t. $B(z, \varepsilon) \subset D$(by full dimensionality). Consider

$$u = z + \varepsilon \frac{T(z)}{\| T(z) \|} \in B(z, \varepsilon).$$

Since z solves (VIP), it implies $0 \geq \langle T(z), u - z \rangle$. However

$$\langle T(z), z + \varepsilon \frac{T(z)}{\| T(z) \|} - z \rangle = \varepsilon \| T(z) \| > 0$$

a contradiction. ■

2.1. An antimonotone (VIP)

In currentsubsection, we consider (VIP(T,D)), where operator $T(\cdot)$ is monotone. We call such a problem an antimonotone (VIP). The name comes after the definition of antimonotone operator (see Def 2.1 and (1)). As noticed in previous section, antimonotone (VIP) is nonmonotone too.

Lemma 2.2. *Let $T(\cdot)$ be a strictly monotone operator, $T(x) \neq 0$ for all $x \in D$; D be a convex subset in $\mathbb{R}^n$. Then if z solves (VIP) then $z \in extr(D)$ (set of extreme points).*

Proof. We consider the following function at z, a solution of antimonotone (VIP).

$$g(x) = \frac{1}{2} \langle \nabla T(z)(x - z), x - z \rangle + \langle T(z), x \rangle.$$

Due to strictly monotonicity of operator $T(\cdot)$ this function is strictly convex [OR70]. Since z solves (VIP) and $\nabla g(x) = \nabla T(z)(x - z) + T(z)$ we have $z = \arg\max\{\langle \nabla g(z), x \rangle \mid x \in D\}$ and

$$\langle \nabla g(z), x - z \rangle \leq 0, \text{ for all } x \in D.$$

It implies that z is a local maximum of the function $g(\cdot)$ on D.

By contradiction let assume that $z \notin extr(D)$: there are two points $u, v \in B(z, \varepsilon) \cap D$ such that $z = \alpha u + (1 - \alpha) v$ for some $0 < \alpha < 1$

$$\begin{cases} g(u) \leq g(z) \\ g(v) \leq g(z) \end{cases} \Longrightarrow \alpha g(u) + (1 - \alpha) g(v) \leq g(z),$$

On the other hand, by strict convexity of $g(\cdot)$

$$g(z) < \alpha g(u) + (1-\alpha) g(v),$$

that completes the proof. ■

Here, we suggest to iteratively solve the antimonotone (VIP) through a sequence of optimization problems with linear objective functions.

$$y^{k+1} = \arg\max \ \{\langle T(y^k), x\rangle \mid x \in D\}. \tag{5}$$

Let us denote $\delta_k = \langle T(y^k), y^{k+1} - y^k\rangle$.

LEMMA 2.3. *If operator $T(\cdot)$ is monotone and continuous, D is compact and at least one of the next two assumptions is true:*

i)$T(\cdot)$ is continuously differentiable and its Hessian $\nabla T(x)$ is symmetric for all $x \in D$;
ii)there is K such that $\forall k \geq K$ the sequence $\{y^{k+1}\}$ fulfills the inequality

$$\langle T(y^{k+1}), y^k\rangle + \delta_k \geq \langle T(y^k), y^k\rangle;$$

then above sequence converges to a solution of antimonotone (VIP).

Proof. i) $\nabla T(x)$ is symmetric for every $x \in D$ if and only if $T(\cdot)$ is integrable; in other words, there exists a Gateaux differentiable function $t(\cdot)$ such that $\nabla t(x) = T(x)$ [OR70]. Monotonicity of $T(\cdot)$ implies that $\nabla T(x)$ is positive semidefinite. Alltogether guarantees convexity of $t(\cdot)$.

On the other hand, according to iterative scheme, we have

$$\langle T(y^k), x - y^{k+1}\rangle \leq 0, \text{ for all } x \in D, \tag{6}$$

that implies $\delta_k = \langle T(y^k), y^{k+1} - y^k\rangle \geq 0$ since $y^k \in D$. Due to convexity of $t(\cdot)$ one gets an increasing numerical sequence $\{t(y^k)\}$ which is bounded by above by $sup\{t(\cdot) \mid D\}(< +\infty$ by Weierstrass's theorem). So, existence of $\lim_{k\to\infty} t(y^k)$ together with continuity of $t(\cdot)$ imply that sequence $\{y^k\}$ has an accumulation point $z = \lim_{k\to\infty} y^k$. Then taking the limit in (6), we have proven that z solves antimonotone (VIP).

ii) From the assumption and monotonocity of $T(\cdot)$, we have the two following inequalities for $k \geq K$:

$$\begin{aligned} \langle T(y^k), y^{k+1} - y^k\rangle + \langle T(y^{k+1}) - T(y^k), y^k\rangle \geq 0; \\ \langle T(y^{k+1} - T(y^k), y^{k+1} - y^k\rangle \geq 0. \end{aligned} \tag{7}$$

Adding them we arrive at $\langle T(y^{k+1}), y^{k+1}\rangle \geq \langle T(y^k), y^k\rangle$ for all k. By continuity of $T(\cdot)$ one gets convergence proof like in case i). ■

Remark. Equivalently, condition ii) in lemma 2.3 above could be written as

$$\langle T(y^k), T(y^k) - y^{k+1}\rangle - \langle y^k, T(y^{k+1}) - y^k\rangle \;\leq\; \| \, T(y^k) - y^k \, \|^2$$

which bears a more algebraic meaning than the operational one used.

3. FROM NONMONOTONE (VIP) TO (PCMP): A BUNDLE TYPE APPROACH

Let D be a nonempty, compact, and convex subset of $\mathbb{R}^n$ and $T : \mathbb{R}^n \to \mathbb{R}^n$ be a continuous operator. This section deals with the variational inequality problem introduced in previous section, $VIP(T,D)$ i.e. to find a vector $z \in D$ such that

$$\langle T(z), x - z\rangle \leq 0, \quad \text{for all } x \in D. \tag{VIP(T,D)}$$

A function $F(.)$ is said piecewise convex if it is defined from convex functions f_i through $F(x) = \min_{m \in M}\{f_m(x)\}$, then the nonconvex and nonsmooth *piecewise convex maximization problem* (PCMP) on D is:

$$\begin{cases} \text{maximize } \; F(x), \\ \text{subject to } x \in D \end{cases} \tag{$PCMP$}$$

We borrow from bundle method in convex programming, the idea of combining descent and proximal point linear approximation in the same framework; for sake of storage capacity, the model kept in the bundle is usually limited to a fixed horizon while the number of proximal points computed could be much greater so that the indices defining the number of elements in the approximation, say l, and the number of steps, say k, are not related.

We aim at giving a sketch for an algorithm solving VIP(T,D), extending bundle idea from linear to convex approximation.

Given a starting point $y^0 \in D$ and some index set $M_0 = \{0\}$. At iteration y^k $(k = 0, 1, 2, ..)$ information $(T(y^k), \nabla T(y^k))$ collected so far, is used to build up a model of some implicit function connected to $VIP(T,D)$. By abuse of notation, we use $\nabla(\cdot)$ for both gradient and Hessian of function and vector function respectively.

Now we construct some function $\tau_l(.)$ such that its gradient equals T at y^k

$$\tau_l(x) = \frac{1}{2}\langle \nabla T(y^k)x, x\rangle + \langle T(y^k) - \nabla T(y^k)y^k, x\rangle, \tag{8}$$

$$\begin{aligned} \nabla\tau_l(x) &= T(y^k) + \nabla T(y^k)(x - y^k), \\ \nabla\tau_l(y^k) &= T(y^k). \end{aligned}$$

Remark. If $y^\star$ is a local maximum of $\tau_l(x)$ over D then $y^\star$ solves VIP(T,D).

It is well known that there are two positive semidefinite matrices $\nabla T(y^k)^+$ and $\nabla T(y^k)^-$ such that

$$\nabla T(y^k) = \nabla T(y^k)^+ - \nabla T(y^k)^-. \tag{9}$$

so that function $\tau_l(x)$ could be written as d.c. (difference of two convex) function

$$\tau_l(x) = \sigma_l(x) - \rho_l(x), \tag{10}$$

where

$$\begin{aligned}
\sigma_l(x) &= \frac{1}{2}\langle \nabla T(y^k)^+ x, x\rangle + \langle T(y^k) - \nabla T(y^k) y^k, x\rangle,\\
\rho_l(x) &= \frac{1}{2}\langle \nabla T(y^k)^- x, x\rangle.
\end{aligned}$$

Due to convexity of $\rho_l(x), \sigma_l(x)$, convex tangent approximation by above and concave tangent approximation by below of $\tau_l(x)$ at y^k are respectively:

$$f_l(x) = \sigma_l(x) - [\rho_l(y^k) + \langle \nabla \rho_l(y^k), x - y^k\rangle] \tag{11}$$

$$g_l(x) = [\sigma_l(x) + \langle \nabla \sigma_l(y^k), x - y^k\rangle] - \rho_l(x) \tag{12}$$

since

$$\begin{aligned}
f_l(x) &\geq \tau_l(x) \geq g_l(x), \text{ for all } x \in \mathbb{R}^n,\\
f_l(y^k) &= \tau_l(y^k) = g_l(y^k).
\end{aligned}$$

Now, we could define a piecewise convex (resp. concave) function to improve model for operator T at iteration $l \in M_l = \{0, 1, 2, ..., l\}$:

$$\begin{aligned}
F_l(x) &= \min\{f_j(x) \mid j \in M_l\}\\
G_l(x) &= \max\{g_j(x) \mid j \in M_l\}
\end{aligned}$$

Obviously neither $F_l(\cdot)$ nor $G_l(\cdot)$ exactly matches desired implicit function but in order to master this lack of information, one enriches the model by finding one more point y^{k+1} and by possibly including one more function $f_{l+1}(\cdot)$ or $g_{l+1}(\cdot)$. To make things simpler, we assign same indices to convex and concave modelling improvement but they could differ according to whether next point improve the model or not. So, we obtain the following **iteration step** $y^k \longmapsto y^{k+1}$.

Algorithm 1 (PCMP for VIP(T,D)).

1. **for** k from 0 while $y^k \notin \arg\max\{F_l(x) \mid x \in D\}$
2. Construct piecewise concave function at y^k: $G_l(x) = \max\{G_{l-1}(x), g_l(x)\}$
3. Compute $u^k = \arg\max\{G_l(x) \mid x \in D\}$
4. Construct piecewise convex function at u^k: $F_l(x) = \min\{F_{l-1}(x), f_l(x)\}$
5. Compute $y^{k+1} = \arg\max\{F_l(x) \mid x \in D\}$
6. **if** $F_l(x) \neq F_{l-1}(x)$
7. **then** $M_{l+1} := M_l \cup \{l+1\}$
8. **else** $M_{l+1} := M_l$
9. **endfor**

First, we solve piecewise concave maximization (an easy problem) at current point then we improve piecewise convex model at solution of previous problem to get next point. Next section delivers an algorithm to solve piecewise convex maximization problem at each iteration, therefore we only have to prove some convergence results.

Lemma 3.1. *The numerical sequence $\{F_l(y^{k+1})\}$ has a limit.*

Proof. First, we prove that $\{F_l(y^{k+1})\}$ is decreasing.
By definition of the function $F_l(\cdot)$ we have

$$F_l(y^{k+1}) = \min\{f_j(y^{k+1}) \mid j \in M_l\} = \min\{F_{l-1}(y^{k+1}), f_l(y^{k+1})\} \leq F_{l-1}(y^{k+1}).$$

Since $F_{l-1}(x) \leq F_{l-1}(y^k)$ for all $x \in D$, as a particular case we have

$$F_{l-1}(y^{k+1}) \leq F_{l-1}(y^k)$$

and therefore $F_l(y^{k+1}) \leq F_{l-1}(y^k)$.

Then we prove that this sequence is bounded by below. From definition of convex tangent approximation

$$\begin{aligned} & f_j(x) \geq \tau_j(x) \text{ for all } x \in \mathbb{R}^n, \quad j \in M_l \\ & \text{and} \\ & F_l(x) = \min\{f_j(x) \mid j \in M_l\} \geq \min\{\tau_j(x) \mid j \in M_l\} \end{aligned} \tag{13}$$

By the Weierstrass's theorem the continuous function $\min\{\tau_j(x) \mid j \in M_l\}$ has a minimum over compact set D. Therefore

$$F_l(y^{k+1}) \geq \min\{\min\{\tau_j(x) \mid j \in M_l\} \mid x \in D\} \tag{14}$$

Hence the convergence of sequence $F_l(y^{k+1})$. ■

Corollary 3.1. *Sequence $\{y^k\}$ has an accumulation point.*

Proof. Since the function $F_l(\cdot)$ is continuous and D is compact, lemma 3.1 implies convergence of the sequence $\{y^k\}$. ■

Proposition 3.1. *A stopping criterion for algorithm 1 is, if $\tau_{l+1}(y^{k+1}) > F_l(y^{k+1})$ then $y^{k+1} = argmax\{F_{l+1}(x) \mid x \in D\}$*

Proof. By assumption of the proposition and by definition of function $f_{l+1}(\cdot)$

$$f_{l+1}(y^{k+1}) \geq \tau_{l+1}(y^{k+1}) > F_l(y^{k+1})$$

holds. Therefore

$$F_{l+1}(y^{k+1}) = \min\{F_l(y^{k+1}), f_{l+1}(y^{k+1})\} = F_l(y^{k+1}) \geq F_l(x) \text{ for all } x \in D,$$

where last inequality comes from y^{k+1} being a maximum of $F_l(x)$. On the other hand,

$$F_l(x) \geq \min\{F_l(x), f_{l+1}(x)\} = F_{l+1}(x) \text{ for all } x \in D.$$

Last two inequalities prove the result

$$F_{l+1}(y^{k+1}) \geq F_{l+1}(x) \text{ for all } x \in D.$$

■

Proposition 3.2. *If at some accumulation point y^{k+1} we have $I(y^{k+1}) = \{l\}$ then y^{k+1} solves VIP(T,D).*

Proof. The definition of the set of active functions at y^{k+1}, $I(y^{k+1}) = \{l\}$ implies $F_l(y^{k+1}) = f_l(y^{k+1})$ and $F_l(y^{k+1}) < f_j(y^{k+1})$ for all $j \in M_l \setminus \{l\}$. So, there exists some neighborhood of y^{k+1} of radius $\delta > 0$ such that for all $x \in B(y^{k+1}, \delta) \cap D$ we have $F_l(x) < f_j(x)$ for $j \in M_l \setminus \{l\}$ and therefore $F_l(x) = f_l(x)$. On the other hand, $y^{k+1} = \arg\max\{F_l(x) \mid x \in D\}$ implies $F_l(x) \leq F_l(y^{k+1})$ for all $x \in D \cap B(y^{k+1}, \delta)$ that means by optimality condition $\langle \nabla F_l(y^{k+1}), x - y^{k+1} \rangle \leq 0$ for all $x \in D \cap B(y^{k+1}, \delta)$.

Since y^{k+1} is an accumulation point, we have $y^k = y^{k+1}$ and $\langle \nabla F_l(y^k), x - y^k \rangle \leq 0$ for all $x \in D \cap B(y^k, \delta)$. It is easy to see that $\langle \nabla F_l(y^k), x - y^k \rangle \leq 0$ for all $x \in D$ since for $w \notin B(y^k, \delta)$ (in other words $\|w - y^k\| > \delta$) one considers $v = y^k + \frac{\delta}{\|w-y^k\|}(w - y^k) \in D$ and $0 \geq \langle \nabla F_l(y^k), v - y^k \rangle = \frac{\delta}{\|w-y^k\|} \langle \nabla F_l(y^k), w - y^k \rangle$.

Now, reminding that $F_l(y^k) = \tau_l(y^k)$ and $\nabla \tau_l(y^k) = T(y^k)$, we can conclude $\langle T(y^k), x - y^k \rangle \leq 0$ for all $x \in D$. $\blacksquare$

4. A PIECEWISE CONVEX MAXIMIZATION (PCMP) ALGORITHM

4.1. optimality conditions

First, we recall from previous section, the definition of (PCMP) for this section to be self-contained; piecewise convex maximization problem (PCMP) is:

$$\begin{cases} \text{maximize } \; F(x), \\ \text{subject to } x \in D \end{cases} \tag{PCMP}$$

where $D \in \mathbb{R}^n$ is a convex and compact set and $F(x) = \min_{m \in M}\{f_m(x)\}$ is defined from convex functions f_i. The purpose of this section is to fully describe a practical algorithm to handle this problem and to give first computational experiments.

We will use as further notations, $clco(D)$ as closure of convex hull of set D and :

$$\begin{aligned} I(z) &= \{i \in M \;/\; f_i(z) = F(z)\}, \\ D_k(z) &= D \cap \{x \mid f_j(x) > F(z) \text{ for all } j \in M \setminus \{k\}\} \end{aligned}$$

for respectively, set of active functions at z, and a special subdomain. In a previous article [Tse00], necessary and sufficient conditions have been derived

PROPOSITION 4.1. *If $z \in D$ is a global maximum of (PCMP) then for all $k \in I(z)$*

$$\partial f_k(y) \bigcap N(D_k(z), y) \neq \emptyset \quad \textit{for all} \quad y \textit{ s.t. } f_k(y) = F(z). \tag{gN}$$

DEFINITION 4.1. F is said **regular** at $z \in D$ if there exists $k \in I(z)$ and $v \in \mathbb{R}^n$ such that $f_k(v) < f_k(z)$.

PROPOSITION 4.2. *If F is regular at $z \in D$ then a sufficient condition for z to be a global maximum for (PCMP) is:*

$$\partial f_k(y) \bigcap N(clco(D_k(z)), y) \neq \emptyset \textit{ for all } y \textit{ s.t. } f_k(y) = F(z) \tag{gS}$$

Based on sufficient condition (gS), a preliminary algorithm was sketched:

ALGORITHM 2 (**PCMP**(x)).

1. Let z be a local maximum of (PCMP) with starting point x.
2. Construct $I(z)$ and choose $s \in I(z)$.
3. Approximate $D_s(z)$ by polytope $\Phi = \{x \in \mathbb{R}^n \mid Px \leq p\}$, where $D_s(z) \subset \Phi \subset D$.

4. **for** l=1 **to** maxiteration
5. y = random point on level set $f_s(y) = F(z)$;
6. $u = \arg\max\{\langle \nabla f_s(y), x\rangle / x \in \Phi\}$; /* linearized problem */
7. **if** $\langle \nabla f_s(y), u - y\rangle > 0$ and $u \in D_s(z)$
8. **then** $x := u$; goto 1; /* better point */
9. **else if** $u \notin D_s(z)$
10. **then** $\Phi := \Phi \bigcap \{x \mid \langle d, x\rangle \leq n, \}$ /* add cutting plane */
11. **endfor**

the following sections will detail this raw algorithm into a runnable code as well as first computational experiments. We consider, in the sequel, only quadratic functions to provide linear gradient and analytic solutions for simple subproblems; then we could use standard solvers in experiments.

$$
\begin{aligned}
f_i(x) &= \frac{1}{2}\langle Q_i x, x\rangle + \langle l_i, x\rangle + \gamma_i \\
Q_i &= [Q_i]^\top, Q_i > 0 \text{ (symmetric positive definite)}
\end{aligned}
$$

4.2. local maximum

This section is devoted to one of the important and difficult problems in nonconvex optimization; local maximum search is known to be in NP-hard class in convex maximization [PS88]. In the literature on global optimization field, it could be oftenly read *let a local maximum be given...* but in practice, an efficient algorithm should handle this assumption since it occurs in inner loop of global maximum search. Here, we emphasize a practical algorithm to address this important issue.

First, let us enhance notations with

$$
\begin{aligned}
\mathcal{L}_f(\alpha) &= \{x \mid f(x) \leq \alpha\} \\
J(z) &= \{j \in M \mid f_j(z) > F(z)\}
\end{aligned}
$$

for Lebesgue set of f at α and Lebesgue related index set. We refine $J(z)$ into $J'(z) = \{j \in M \mid f_j(z) > F(z), \ \ \mathcal{L}_{f_j}(F(z)) \neq \emptyset\}$ and accordingly M as $M' = I \cup J'$; in practice, it may happen that $J' \neq J$ at a given z due to rounding errors, requiring above refinement in local search algorithm or else a premature emptiness stopping criterion is faulty detected.

Let x^k be a feasible point in (PCMP), by definition of $F(.)$ both $I(x^k) \neq \emptyset$ and $x^k \notin \mathcal{L}_{f_j}(F(x^k))$ for all $j \in J(x^k)$.

Remark. As initial feasible point, we choose maximum over all quadratic minimizations, $x^0 = \max_{m\in M}\{\min_x\{f_m(x) \mid x \in D\}$.

In the remainder of the section, we will drop arguments for sake of concineness and write I, J instead of $I(x^k), J(x^k)$.

We introduce a set of polytopes P^k related to current point x^k:

$$P^k = \left\{ x \in \mathbb{R}^n \mid \begin{array}{ll} \langle \nabla f_j(x^k), x \rangle \geq \langle \nabla f_j(x^k), v^j \rangle, & j \in J \\ \langle \nabla f_i(x^k), x \rangle \geq \langle \nabla f_i(x^k), x^k \rangle, & i \in I \end{array} \right\}$$

where $v^j = \arg\max\{\langle \nabla f_j(x^k), x \rangle \mid x \in \mathcal{L}_{f_j}(F(x^k))\}, \quad j \in J$ could be analytically solved under quadratic assumption:

$$\begin{aligned} &\arg\max\{\langle d, x \rangle \ \text{ s.t. } \ \frac{1}{2}\langle Qx, x \rangle + \langle l, x \rangle + \gamma \leq \xi\} \\ &= \ Q^{-1}(d\sqrt{\frac{\langle Q^{-1}l, l \rangle + 2(\xi - \gamma)}{\langle Q^{-1}d, d \rangle}} - l) \end{aligned} \tag{15}$$

where d, Q, l, γ are appropriately associated with functions f_j and $\xi = F(x^k)$

Algorithm 3 (**Local Maximum**(D, M)**).**

1. Let $x^k \in D$ be given
2. **for all** $m \in M'$
3. $\quad w^m = \arg\max\langle \nabla f_m(x^k), x \rangle$ s.t. $x \in D \cap P^k$
4. **endforall**
5. $w^r = \arg\max\{\min_{i \in M} f_i(w^m) \mid m \in M'\}$
6. **if** $\| w^r - x^k \| \leq \epsilon$
7. $\quad$ **then** stop /* local maximum found */
8. $\quad$ **else** $k = k + 1; x^k = w^r$; goto 2;

Remark. If we strictly follow sufficient optimality condition, we need examine only active function i.e. $m \in I$ instead of $m \in M'$; however, we are looking for a point in complement of Lebesgue set of F, therefore a much symmetrized behavior comes from selecting all functions in $M' = I \cup J$ instead. We could notice too, in step 3, that while w^r is extracted from a $M \times M'$ array of values, it is actually found through a 1 dimensional loop along all objectives $f_i, \ i \in M$.

Proposition 4.3. *Numerical sequence $\{F(x^k)\}$ is non decreasing and $\lim_{k \to \infty} x^k$ is a local maximum of (PCMP).*

Proof. since $w^r = \arg\max\{\min_{i\in M} f_i(w^m) \mid m \in M'\}$, we have $w^r \in D \cap P^k$. By definition of polytopes P^k and using convexity:

$$\begin{aligned} 0 &\leq \langle \nabla f_i(x^k), w^r - x^k \rangle \leq f_i(w^r) - f_i(x^k), \;\; i \in I \\ 0 &\leq \langle \nabla f_j(v^j), w^r - v^j \rangle = \lambda \langle \nabla f_j(x^k), w^r - v^j \rangle \leq \lambda (f_j(w^r) - f_j(v^j)), \;\; j \in J \end{aligned}$$

where $\lambda > 0$ is directly derived from (15). From (15) and definition of I, we arrive at

$$\begin{aligned} F(x^k) &= f_i(x^k) \leq f_i(w^r), \;\; i \in I \\ F(x^k) &= f_j(v^j) \leq f_j(w^r), \;\; j \in J \end{aligned}$$

In other words, we get $F(x^k) \leq F(x^{k+1})$.

On the other hand, for all $m \in M', F(x^{k+1}) \geq F(w^m) \geq F(x^k)$, so let assume $z = x^k$ is an accumulation point for some k; then, for some $\epsilon > 0$ for all d such that $z + \epsilon d \in D \cap P^k$, we have $F(z) = F(w^m) \geq F(z + \epsilon d)$ since $w^i = \arg\max\{\langle \nabla f_i(z), x \rangle \mid P^k\}$, hence the result. ■

Remark. In above proof, stopping criterion comes from emptiness of polytope P^k; however, in practice, we observed slow speed of convergence and tailing off effect close to local maximum. Therefore, we adopt a multiresolution scheme and early force emtyness of P^k for a given accuracy $\varepsilon = 10^l \times \epsilon$ (a multiple of actual tolerance ϵ); then we refine $\varepsilon = \varepsilon/10$ until $\varepsilon \leq \epsilon$. Under this multiresolution scheme, we get much faster convergence without tailing off effect on examples reported below but this practical rule of the thumb requires a thorough sensibility analysis to prove speed of convergence from coarse to finest resolution. Under this scheme, polytopes P^k_ε are introduced as inner approximation of above P^k.

$$P^k_\varepsilon = \left\{ x \in \mathbb{R}^n \mid \begin{array}{l} \langle \nabla f_j(x^k), x \rangle \geq \langle \nabla f_j(x^k), v^j \rangle + \varepsilon, \;\; j \in J \\ \langle \nabla f_i(x^k), x \rangle \geq \langle \nabla f_i(x^k), x^k \rangle + \varepsilon, \;\; i \in I \end{array} \right\}$$

4.3. separation: an intersection graph

Given a local maximum z, in order to improve best known solution we have to look for a point in $\mathrm{clco}(D_m(z))$, a difficult problem whose construction was bypassed by recoursing to Caratheodory's existence theorem. In this section, we borrow from linear programming, the well-known equivalence **separation** $\equiv$ **optimization**, to notice that whenever 2 functions f_i, f_j are disjoint then it's likely to retrieve a $\mathrm{clco}(D_m(z))$ point inbetween. Once more, while from theoretical viewpoint, $\mathrm{clco}(D_m(z))$ is searched for, for only one $m \in M$, a better behavior was observed under isotropic retrieval $\cap_{m\in M}\mathrm{clco}(D_m(z))$. Here, we address intersection graph construction, namely we build a symmetric graph $\mathcal{G}(z) = (V, E(z))$ having

$V = \{f_i, \quad i \in M\}$ as vertices and $E(z) = \{(f_i, f_j) \mid D \cap \mathcal{L}_{f_i}(F(z)) \cap \mathcal{L}_{f_j}(F(z)) \neq \emptyset\}$ as edges. To decide whether functions f_i and f_j are separated, we turn separation into both optimization problems:

$$\min f_i(x) \ \text{ s.t. } \ x \in D \cap \mathcal{L}_{f_j}(F(z))$$
$$\min f_j(x) \ \text{ s.t. } \ x \in D \cap \mathcal{L}_{f_i}(F(z))$$

Once more, we introduce a set of polytopes that outer approximate Lebesgue sets $\mathcal{L}_{f_i}(F(z))$ and $\mathcal{L}_{f_j}(F(z))$ by piecewise linear hyperplanes at some points on level sets. We have $f_i(\bar{x_i}^k) + \langle \nabla f_i(\bar{x_i}^k), x - \bar{x_i}^k \rangle \leq f_i(x) \leq F(z)$ using convexity and Lebesgue set definition; then selecting $\bar{x_i}^k$ on level set, $f_i(\bar{x_i}^k) = F(z)$, we approximate Lebesgue set as announced: $P_i^k = \cap_{k\geq 0}\{\langle \nabla f_i(\bar{x_i}^k), x - \bar{x_i}^k \rangle \leq 0\}$, with initial condition $P_i^0 = \mathbb{R}^n$ to be consistent with algorithm below. We introduce, in the same fashion: $P_j^k = \cap_{k\geq 0}\{\langle \nabla f_j(\bar{x_j}^k), x - \bar{x_j}^k \rangle \leq 0\}$ and $P_j^0 = \mathbb{R}^n$.

Algorithm 4 (**Intersection edge**(f_i, f_j)).

1. $k = 0$;
2. **forever do**
3. $x_i^k = \arg\min\{f_i(x) \mid x \in D \cap P_j^k\}$;
4. $x_j^k = \arg\min\{f_j(x) \mid x \in D \cap P_i^k\}$;
5. **if** $(f_i(x_i^k) \leq F(z)$ and $f_j(x_i^k) \leq F(z))$ or $(f_i(x_j^k) \leq F(z)$ and $f_j(x_j^k) \leq F(z))$
6. **then return** edge(f_i, f_j);
7. **if** $(f_i(x_i^k) > F(z)$ and $f_j(x_i^k) > F(z))$ or $(f_i(x_j^k) > F(z)$ and $f_j(x_j^k) > F(z))$
8. **then return** no edge(f_i, f_j);
9. **enddo**

The correctness of this algorithm follows $P_i^k \supseteq P_i^{k+1}$ and $P_j^k \supseteq P_j^{k+1}$ so that either condition on edge(f_i, f_j) is fulfilled for some k.

4.4. inner clco approximation

For all separated objectives in $\mathcal{G}(z)$, we could guess whether a point belongs to $\cap_{m\in M}\text{clco}(D_m(z))$ in the following way; let consider f_i, f_j be separated in D with $\bar{x_i}, \bar{x_j}$ as points on respective level sets $f_i(\bar{x_i}) = F(z), f_j(\bar{x_j}) = F(z)$ then $x = \frac{1}{2}(\bar{x_i} + \bar{x_j})$ is likely to belong to $\cap_{m\in M}\text{clco}(D_m(z))$ or to be a good starting point to look for a local maximum over all objectives but i and j; let M_{ij} denote $M\backslash\{i,j\}$, $D_{ij} = D \cap \{\langle \nabla f_j(\bar{x_j}), x - \bar{x_j} \rangle \geq 0, \langle \nabla f_j(\bar{x_j}), x - \bar{x_j} \rangle \geq 0\}$, then solving **Local Maximum**(D_{ij}, M_{ij}) ends up with either a better point in $D_{ij} \subset \cap_{m\in M}\text{clco}(D_m(z))$ or no better point in D_{ij} which does not mean $\cap_{m\in M}\text{clco}(D_m(z))$ is empty.

4.5. outer clco approximation

In preliminary algorithm PCMP(x) step 4, we borrow from global optimization a random generation of points on level sets; while in global maximization it affords to cut current point, in PCMP case it amounts to approximate $\cap_{m\in M}\text{clco}(D_m(z))$ from outside (unless the random point belongs to clco in which case we could look for a local maximum like in inner clco approximation scheme). In multiknapsack maximization, a *geodesic partitioning property* [FT00] allows us to reduce random generation to some representative cone and yields an efficient algorithm; in PCMP case, this appealing technique has to be carried over $\cap_{m\in M}\text{clco}(D_m(z))$, an open problem.

As a first attempt towards this goal, we substitute to random point generation, selection of tangent point to objective f_s along each constraint direction and then solve linearized problem at that deterministic point y. Starting with polytope $\Phi = D$, we refine it through cutting planes; let us compute $r = arg\min_j\{f_j(u) \mid j \in M\}$ and index set of **active constraints** in current Φ at u

$$K(u) = \{l \mid [Pu]_l = p_l\}$$

where P (resp. $P(u)$) denotes the matrix of constraints (resp. active constraints at u) of polytope Φ; under full dimensionality assumption, $[P(u)]^{-1}$ definitely exists. Let Y the set (columnwise) of points on level set $f_r(\cdot)$ intersected by active cone, namely

$$Y = u \otimes e^\top - [P(u)]^{-1}\alpha^r,$$

where $\otimes$ denotes kronecker product and $\alpha^r \in \mathbb{R}^n_+$ solve the quadratic equations for every column vector y^i of Y

$$f_r(y^i(\alpha)) = F(z^k).$$

Then vector d, found as solution of the linear system $Yd = ne$ yields a new cut for polytope Φ, as well known in global optimization field; notice that right handside introduces a normalizing factor to avoid tailing off effects since it is usual to observe such effect in similar algorithms [FT00], [ST98]).

4.6. computational experiments

We report software experiments on examples designed for optimality conditions; in order to thoroughly test algorithm, we introduce some variants like discarding useless objective, extending domain D...

EXAMPLE 4.1.

$$\begin{aligned} f_1(x) &= x_1^2 + (x_2+4)^2 - 36, \\ f_2(x) &= (x_1+8)^2 + (x_2-3)^2 - 36, \\ f_3(x) &= x_1^2 + (x_2-8)^2 - 16, \end{aligned}$$

$$f_4(x) = (x_1 - 8)^2 + (x_2 - 3)^2 - 53,$$
$$f_5(x) = (x_1 - 10)^2 + (x_2 + 10)^2 - 4$$

subject to

$$D = \{x \in \mathbb{R}^2 \mid -4 \le x_1 \le 10, \ \ -6 \le x_2 \le 8, \ \ x_1 - x_2 \le 10\}.$$

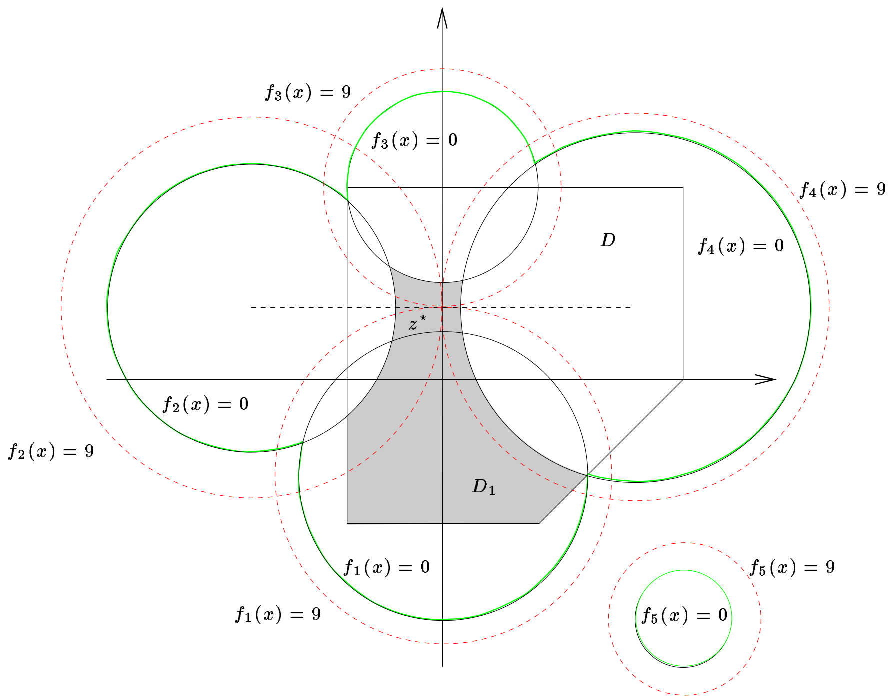

FIG. 3. a simple example **??**

EXAMPLE 4.2.

$$f_1(x) = x_1^2 + (x_2 + 2)^2 - 9,$$
$$f_2(x) = 9(x_1 + 3)^2 + 4x_2^2 - 36,$$
$$f_3(x) = (x_1 + 1)^2 + (x_2 - 4)^2 - 4,$$
$$f_4(x) = \frac{1}{9}(x_1 - 3)^2 + \frac{1}{36}(x_2 - 4)^2 - 1,$$

$$f_5(x) = (x_1 - 5)^2 + (x_2 + 5)^2 - 1$$

subject to box constraint:

$$D = \{x \in \mathbb{R}^2 \mid -2 \le x_1 \le 5, \;\; -3 \le x_2 \le 4\}.$$

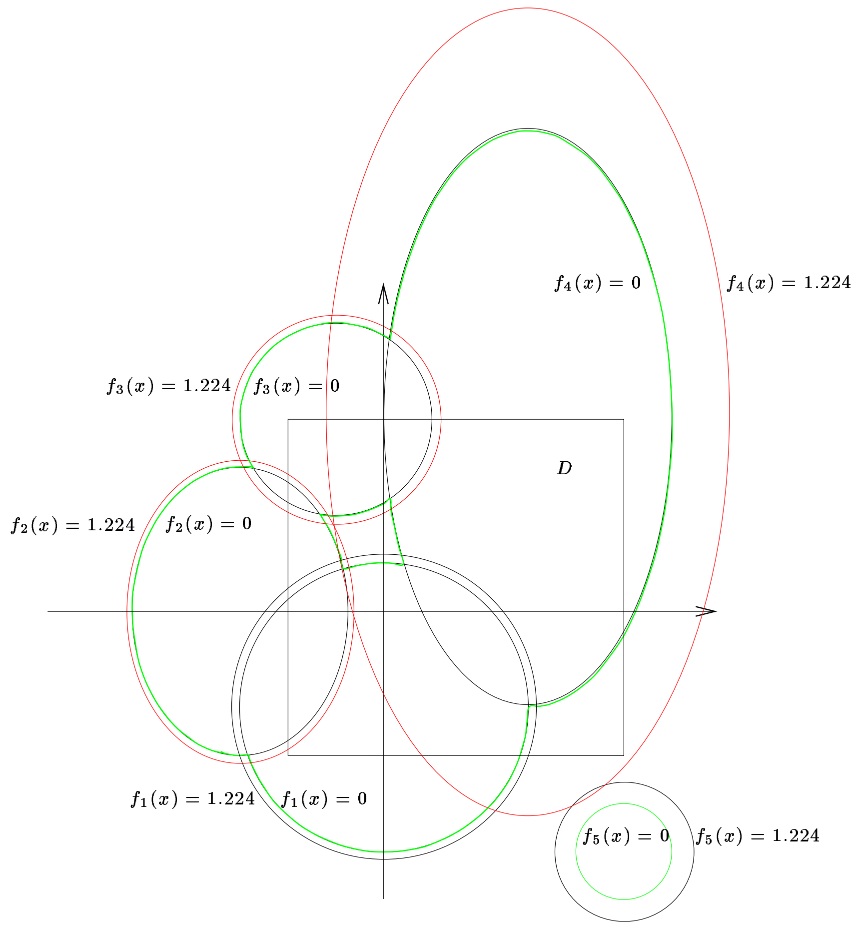

FIG. 4. a non trivial example 4.2

Variant to fully test intersection graph:

EXAMPLE 4.3.

$$
\begin{aligned}
f_1(x) &= x_1^2 + (x_2+4)^2 - 36, \\
f_2(x) &= (x_1+8)^2 + (x_2-3)^2 - 36, \\
f_3(x) &= x_1^2 + (x_2-8)^2 - 16, \\
f_4(x) &= (x_1-8)^2 + (x_2-3)^2 - 53, \\
f_5(x) &= (x_1-10)^2 + (x_2+10)^2 - 4 \\
f_6(x) &= (x_1+5)^2 + (x_2+4)^2 - 1
\end{aligned}
$$

subject to

$$D = \{x \in \mathbb{R}^2 \mid -4 \leq x_1 \leq 10, \;\; -6 \leq x_2 \leq 8, \;\; x_1 - x_2 \leq 10\}.$$

Variants to handle non regularity; a problem happens to be non regular (see definition 4.1) as soon as one center of the ellipses lay on the boundary of D while the remaining ellipses are not active at this center. In local maximum search, it leads to a null gradient side effect resulting in an erratic (and slow) trajectory towards accumulation point. In order to circumvent this bad effect in practice, we escape from local search and directly apply inner clco approximation; it considerably speeds up algorithm since a deep clco point is quickly retrieved.

EXAMPLE 4.4.

$$
\begin{aligned}
f_1(x) &= x_1^2 + (x_2+4)^2 - 36, \\
f_2(x) &= (x_1+8)^2 + (x_2-3)^2 - 36, \\
f_3(x) &= x_1^2 + (x_2-8)^2 - 16, \\
f_4(x) &= (x_1-8)^2 + (x_2-3)^2 - 53,
\end{aligned}
$$

subject to

$$D = \{x \in \mathbb{R}^2 \mid -4 \leq x_1 \leq 10, \;\; -6 \leq x_2 \leq 8, \;\; x_1 - x_2 \leq 10\}.$$

EXAMPLE 4.5.

$$
\begin{aligned}
f_1(x) &= x_1^2 + (x_2+2)^2 - 9, \\
f_2(x) &= 9(x_1+3)^2 + 4x_2^2 - 36, \\
f_3(x) &= (x_1+1)^2 + (x_2-4)^2 - 4, \\
f_4(x) &= \frac{1}{9}(x_1-3)^2 + \frac{1}{36}(x_2-4)^2 - 1,
\end{aligned}
$$

subject to box constraint:

$$D = \{x \in \mathbb{R}^2 \mid -2 \le x_1 \le 5, \;\; -3 \le x_2 \le 4\}.$$

Variants lifted to 3 dimensions. The same examples are lifted to a 3D box and size of box was chosen either symmetric or non symmetric to measure how algorithm could escape from very good point lifted from 2D case whenever box is large enough.

EXAMPLE 4.6.

$$\begin{aligned} f_1(x) &= x_1^2 + (x_2+4)^2 + x_3^2 - 36, \\ f_2(x) &= (x_1+8)^2 + (x_2-3)^2 + x_3^2 - 36, \\ f_3(x) &= x_1^2 + (x_2-8)^2 + x_3^2 - 16, \\ f_4(x) &= (x_1-8)^2 + (x_2-3)^2 + x_3^2 - 53, \\ f_5(x) &= (x_1-10)^2 + (x_2+10)^2 + x_3^2 - 4 \end{aligned}$$

subject to

$$D = \{x \in \mathbb{R}^2 \mid -4 \le x_1 \le 10, \;\; -6 \le x_2 \le 8, \;\; -2 \le x_3 \le 2, \;\; x_1 - x_2 \le 10\}.$$

EXAMPLE 4.7.

$$\begin{aligned} f_1(x) &= x_1^2 + (x_2+4)^2 + x_3^2 - 36, \\ f_2(x) &= (x_1+8)^2 + (x_2-3)^2 + x_3^2 - 36, \\ f_3(x) &= x_1^2 + (x_2-8)^2 + x_3^2 - 16, \\ f_4(x) &= (x_1-8)^2 + (x_2-3)^2 + x_3^2 - 53, \\ f_5(x) &= (x_1-10)^2 + (x_2+10)^2 + x_3^2 - 4 \end{aligned}$$

subject to

$$D = \{x \in \mathbb{R}^2 \mid -4 \le x_1 \le 10, \;\; -6 \le x_2 \le 8, \;\; -2 \le x_3 \le 4, \;\; x_1 - x_2 \le 10\}.$$

EXAMPLE 4.8.

$$f_1(x) = x_1^2 + (x_2+2)^2 + x_3^2 - 9,$$

$$\begin{aligned}
f_2(x) &= 9(x_1+3)^2 + 4x_2^2 + x_3^2 - 36, \\
f_3(x) &= (x_1+1)^2 + (x_2-4)^2 + x_3^2 - 4, \\
f_4(x) &= \frac{1}{9}(x_1-3)^2 + \frac{1}{36}(x_2-4)^2 + x_3^2 - 1, \\
f_5(x) &= (x_1-5)^2 + (x_2+5)^2 + x_3^2 - 1
\end{aligned}$$

subject to box constraint:

$$D = \{x \in \mathbb{R}^2 \mid -2 \leq x_1 \leq 5, \;\; -3 \leq x_2 \leq 4, \;\; -2 \leq x_3 \leq 2\}.$$

Example 4.9.

$$\begin{aligned}
f_1(x) &= x_1^2 + (x_2+2)^2 + x_3^2 - 9, \\
f_2(x) &= 9(x_1+3)^2 + 4x_2^2 + x_3^2 - 36, \\
f_3(x) &= (x_1+1)^2 + (x_2-4)^2 + x_3^2 - 4, \\
f_4(x) &= \frac{1}{9}(x_1-3)^2 + \frac{1}{36}(x_2-4)^2 + x_3^2 - 1, \\
f_5(x) &= (x_1-5)^2 + (x_2+5)^2 + x_3^2 - 1
\end{aligned}$$

subject to box constraint:

$$D = \{x \in \mathbb{R}^2 \mid -2 \leq x_1 \leq 5, \;\; -3 \leq x_2 \leq 4, \;\; -2 \leq x_3 \leq 4\}.$$

All examples were run under digital PWS500 unix workstation using CPLEX solver and C++. In tables below, we split results into 2 parts: one for inner clco approximation only and one for the complete loop through inner then outer clco approximation until outer approximation could not be refined or a maximum of 10 iterations occurs.

Meanings for all columns in tables follow:

CP: #hyperplanes in finest outer clco approximation,

best: best value known on corresponding approximation,

fat: how bad outer clco approximation could be, namely given the minimum value over each objective f_m considered by itself, i.e. $q = \min_{m \in M}\{f_m(\arg\min_y\{f_m(y) \mid y \; in \Phi\}$ and actual best known value $Q = \min_{m \in M}\{f_m(y) \mid y \; in \Phi\}$, this ratio (in percentage) is equal to $100 * (Q-q)/fabs(q)$; on the contrary, a small value means that outer clco approximation is tight,

time: overall (inner+outer CLCO) user time in seconds,

TABLE 1

outer approximation cutting plane: *worse inactive + all objectives*

	inner CLCO				inner+outer CLCO							
ex.	local	inner	graph	best	best	local	inner	graph	outer	CP	fat%	time
4.1	5:476	7:96	48:0	11.8676	11.8676	5:476	7:96	48:0	0:720	47	156	3.737
4.3	6:780	8:64	65:0	11.8671	11.8671	6:780	8:64	65:0	0:1524	49	156	8.045
4.4	4:340	2:64	36:0	11.8677	11.8677	4:340	2:64	36:0	0:960	51	156	4.735
4.2	5:64	5:39	47:0	1.2237	1.2237	5:64	5:39	47:0	0:1630	50	12	7.488
4.5	4:64	2:16	58:0	1.22335	1.22335	4:64	2:16	58:0	0:1216	50	121	5.423
4.6	5:418	7:93	48:0	15.8672	15.8678	5:624	17:279	96:0	0:7962	921	179	784.958
4.7	5:418	7:93	48:0	15.8672	15.8672	5:418	7:93	48:0	0:15445	699	169	31064.289
4.8	5:46	0:18	48:0	5.22369	5.22392	5:57	0:21	69:0	0:8916	654	201	1259.574
4.9	5:46	0:18	48:0	5.22369	16.8056	5:73	0:42	66:0	0:4821	213	425	349.588

all remaining columns: number of quadratic minimizations and number of linear programs solved for related step written as #QPs:#LPs.

For outer clco approximation, we test different strategies for refining current approximation: given $u = \arg\max$ of linearized problem, we select either **worse** objective or **closest** to active objective at u; we also test symmetrization of clco like for inner approximation and generate tangent point along all constraints either for **all** or only for **active** objectives at local maximum z. Table 1 reports worse selection from u and all objectives tangent point generation, table 2 reports worse selection from u and active objectives only and table 3 reports closest to active selection (if any) and active objectives only.

As first concluding remark, we would say that inner clco approximation achieves very good solution in a small amount of effort and that outer clco approximation is not tight and scarcely improves inner clco result; moreover, outer approximation introduces many hyperplanes before stopping according to some tolerance factor and requires further sensibility analysis because we observe, for 3D examples, overdetermined active cone (inducing gaussian elimination with pivoting even under full dimensionality assumption) and non invertible cone (due to so many hyperplanes under rounding error; when it arises, we discard such a numerically unstable hyperplane in practical algorithm though a more sophisticated code would need SVD computation).

4.7. towards an inclusion-exclusion algorithm

Despite the poor outer clco approximation scheme, we suspect that intersection graph endows non convex clco with a very rich structure. We only exploited non adjacency between 2 vertices (separation between 2 objectives) to inner approximate clco; in this section,

TABLE 2

outer approximation cutting plane: ***worse inactive + active objectives***

	inner CLCO				inner+outer CLCO							
ex.	local	inner	graph	best	best	local	inner	graph	outer	CP	fat%	time
4.1	5:476	7:96	48:0	11.8676	11.8676	5:476	7:96	48:0	0:168	23	155	1.220
4.3	6:780	8:64	65:0	11.8671	11.8671	6:780	8:64	65:0	0:168	23	155	2.138
4.4	4:340	2:64	36:0	11.8677	11.8677	4:340	2:64	36:0	0:168	23	155	1.052
4.2	5:64	5:39	47:0	1.2237	1.2237	5:64	5:39	47:0	0:1012	51	122	3.602
4.5	4:64	2:16	58:0	1.22335	1.22335	4:64	2:16	58:0	0:1012	51	122	3.636
4.6	5:418	7:93	48:0	15.8672	15.8672	5:418	7:93	48:0	0:1272	246	176	69.856
4.7	5:418	7:93	48:0	15.8672	15.8672	5:418	7:93	48:0	0:394	155	160	8.868
4.8	5:46	0:18	48:0	5.22369	5.22369	5:46	0:18	48:0	0:386	182	197	13.581
4.9	5:46	0:18	48:0	5.22369	16.8056	5:73	0:42	66:0	0:2671	236	365	272.398

TABLE 3

outer approximation cutting plane: ***closest to active + active objectives***

	inner CLCO				inner+outer CLCO							
ex.	local	inner	graph	best	best	local	inner	graph	outer	CP	fat%	time
4.1	5:476	7:96	48:0	11.8676	11.8676	5:476	7:96	48:0	0:964	74	159	5.709
4.3	6:780	8:64	65:0	11.8671	11.8671	6:780	8:64	65:0	0:1444	79	156	9.065
4.4	4:340	2:64	36:0	11.8677	11.8677	4:340	2:64	36:0	0:964	74	159	5.476
4.2	5:64	5:39	47:0	1.2237	1.2237	5:64	5:39	47:0	0:860	65	120	4.849
4.5	4:64	2:16	58:0	1.22335	1.22335	4:64	2:16	58:0	0:868	66	123	12.001
4.6	5:418	7:93	48:0	15.8672	15.8672	5:418	7:93	48:0	0:1194	517	178	252.815
4.7	5:418	7:93	48:0	15.8672	15.8672	5:418	7:93	48:0	0:88	51	142	1.173
4.8	5:46	0:18	48:0	5.22369	5.22369	5:46	0:18	48:0	0:4688	328	204	1159.106
4.9	5:46	0:18	48:0	5.22369	16.8056	5:73	0:42	66:0	0:813	245	371	60.163

we provide a few arguments to open a possible further study of intersection graph since experiments for 3D cases have shown that it could be difficult to escape from a rather *good* clco situation. In fig. 5 we depicted the situation when we end up with an empty inner approximation (upto tolerance) while outer approximation is far from actual clco extreme points; in left picture, we assume that inner approximation of $\cap_{m\in\{i,j,k\}}\mathrm{clco}(D_m(z))$ has to be augmented with non convex part coming from outer clco approximation and in right picture we are concerned with actual situation in practical algorithm. So next step would be to deal with the following PCMP optimization: let $F_{ij} = \min_{m\in M_{ij}}\{f_m(x)\}$ for non adjacent f_i, f_j in intersection graph

$$\begin{aligned} \text{maximize} \quad & F_{ij}(x) \\ \text{subject to} \quad & x \in D \\ & \langle \nabla f_i(\bar{x}_i), x - \bar{x}_i \rangle \geq 0 \qquad (16) \\ & \langle \nabla f_j(\bar{x}_j), x - \bar{x}_j \rangle \geq 0 \qquad (17) \end{aligned}$$

where as above $M_{ij} = M\backslash\{i,j\}$ and constraints 16 and 17 include some convex part of clco related to closest apart points $\bar{x}_i, \bar{x}_j$ on level sets $f_i(x) = f_j(x) = F(z)$. It could be extended to adjacent f_i, f_j as well, provided closest apart points $\bar{x}_i, \bar{x}_j$ in $\mathcal{L}_{f_i}(F(z)) \cap \mathcal{L}_{f_i}(F(z))$ are *close enough* to a clco extreme point. In more general words, patching convex inner clco approximation with non convex parts from outer clco approximation is a challenging issue around the famous inclusion-exclusion principle on property ψ on sets A, B, C:

$$\psi(A \cup B \cup C) = \psi(A) + \psi(B) + \psi(C) - \psi(A \cap B) - \psi(A \cap C) - \psi(B \cap C) + \psi(A \cap B \cap C)$$

In practical experiments we deal with only 2 sets and sketch above how to deal with 3 sets; as for its general setting with n sets, it remains far beyond the scope of this prospective concluding remarks.

Another minor remark concerns the intersection graph $\mathcal{G}(z)$ itself. It is wortwhile to concentrate on connected components in turn since each corresponding clco is a candidate for improvement; however, no clear understanding of their relative *degree of improvement* is available. In the same way, modular decomposition of $\mathcal{G}(z)$ could help in choosing a *direction* of improvement without any clear understanding of such an impact.

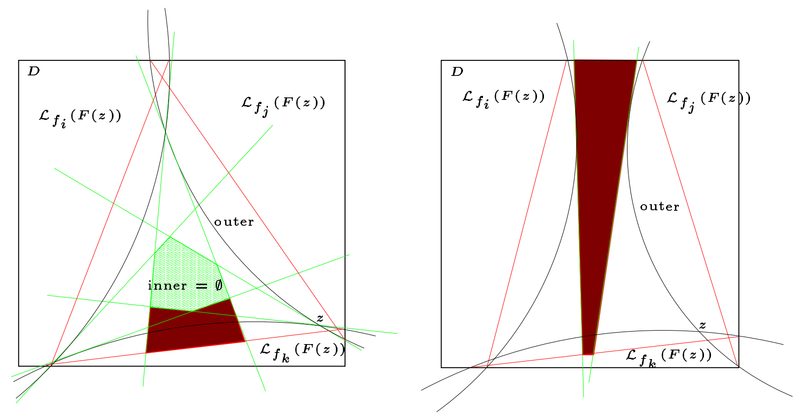

FIG. 5. an inclusion-exclusion clco approximation

5. CONCLUDING REMARKS

In this article, we gave a practical algorithm to solve piecewise convex maximization problems; it has been proven efficient to retrieve optimal solution in $\mathbb{R}^2$ and at the same time it suggested how difficult it could be to escape from a very good local maximum. We introduced the intersection graph between objectives and noticed how it influences the direction of search; it compares favorably with *standard* techniques from global optimization that amount to outer approximate the region of interest through hyperplanes. It opens a field to study more thoroughly:

- the structure of this graph (connected components, modular decomposition,...)
- its connection to the inclusion-exclusion principle applied to convex and non convex parts of the crucial parts in the solution space,
- its relationships with finding the real solutions of a system of non linear equations [Bul00].

To our knowledge, any further improvement will have a great impact in solving variational inequalities and related problems as mentionned in introductory sections.

References

[Ben95] H. P. Benson. *Concave Minimization: Theory, applications and algorithms.* Kluwer Academic, Dordrecht/Boston/London, 1995.

[Bul00] V. P. Bulatov. Numerical method for finding all real roots of systems of nonlinear equations. (in russian). *Zh. Vychisl. Mat. Mat. Fiz.*, 40(3):348–355, 2000.

[FT00] D. Fortin and I. Tsevendorj. Global optimization and multiknapsack: a percolation algorithm. *Rapport de recherche INRIA, France*, (3912):19, April 2000.

[HPT95] R. Horst, P.M. Pardalos, and N. Van Thoai. *Introduction to Global Optimization.* Kluwer Academic, 1995.

[HT90] R. Horst and H. Tuy. *Global Optimization.* Springer-Verlag, 1990.

[HU89] J. B. Hiriart-Urruty. *From Convex Optimization to Nonconvex Optimization, Part I. Necessary and Sufficient Conditions for Global Optimality.* Plenum, New York, 1989.

[HU95] J. B. Hiriart-Urruty. *Conditions for Global Optimality.* Kluwer Academic, Dordrecht/Boston/London, 1995.

[LR96] J. S. Luo, Z. Q.and Pang and D. Ralph. *Mathematical Programs with Equilibrium Constraints.* Cambridge University Press, 1996.

[OR70] J. M. Ortega and W. C. Rheinboldt. *Iterative solution of nonlinear equations in several variables.* Academic Press, 1970.

[PS88] P. M. Pardalos and G. M. Schnitger. Checking local optimality in constrained quadratic problems is np-hard. *Oper. Res. Letters*, 7:33–35, 1988.

[Roc70] R. T. Rockafellar. *Convex Analysis.* Princeton University Press, 1970.

[ST98] A. S. Strekalovsky and I Tsevendorj. Testing the r-strategy for a reverse convex problem. *J. Global Optim.*, 13(1):61–74, 1998.

[Str87] A. S. Strekalovskii. On the global extremum problem. *Soviet Dokl.*, (292):1062–1066, 1987.

[Str93] A. S. Strekalovsky. The search for a global maximum of a convex functional on an admissibleset. *Computational Mathematics and Mathematical Physics*, 33(3):315–328, 1993.

[Str98] A. S. Strekalovsky. Global optimality conditions for nonconvex optimization. *J. Global Optim.*, 12(4):415–434, 1998.

[Tse98a] I. Tsevendorj. On the conditions for global optimality. *Journal of the Mongolian Mathematical Society*, 1(2):58–61, 1998.

[Tse98b] I. Tsevendorj. To a global solution characterization of nonconvex problems. In *11-th Baikal International School-Seminar on Optimization Methods and its applications*, pages 212–215, Irkutsk, Russia, 1998.

[Tse00] I. Tsevendorj. Optimality conditions for piecewise-convex maximization problems. *Rapport de recherche INRIA, France*, (3941):20, May 2000.

Unité de recherche INRIA Lorraine, Technopôle de Nancy-Brabois, Campus scientifique,
615 rue du Jardin Botanique, BP 101, 54600 VILLERS LÈS NANCY
Unité de recherche INRIA Rennes, Irisa, Campus universitaire de Beaulieu, 35042 RENNES Cedex
Unité de recherche INRIA Rhône-Alpes, 655, avenue de l'Europe, 38330 MONTBONNOT ST MARTIN
Unité de recherche INRIA Rocquencourt, Domaine de Voluceau, Rocquencourt, BP 105, 78153 LE CHESNAY Cedex
Unité de recherche INRIA Sophia-Antipolis, 2004 route des Lucioles, BP 93, 06902 SOPHIA-ANTIPOLIS Cedex

Éditeur
INRIA, Domaine de Voluceau, Rocquencourt, BP 105, 78153 LE CHESNAY Cedex (France)
http://www.inria.fr
ISSN 0249-6399